Efficiënt zoeken op het net
Virtuele speurneuze

Niks makkelijker dan surfen op het web: snel even dat adres intikken en voor je het weet, verschijnt de pagina al op je scherm! Maar wat als je geen flauw idee hebt van het adres en je wil toch snel een paar bruikbare documenten bij elkaar sprokkelen? Tja, dan is het zoeken geblazen, en temidden van miljarden webpagina's is dat niet altijd evident...



Kristien koos deze cursus.

Zoekknop



Een ingebouwde zoekknop... handig, dat wel.

Als beginnend surfer lijkt de ingebouwde zoekknop van je *browser* (de vaktaal vind je in het midden van dit nummer) natuurlijk erg uitnodigend! Maar wat gebeurt er nu precies als je die knop indrukt in een browser als Internet Explorer? Aan de linkerzijde opent zich dan een taakvenster met een witte balk waarin je een zoekterm kan typen (die overigens uit verschillende woorden mag bestaan). Vervolgens druk je op de knop ZOEKEN en met wat geluk verschijnen prompt enkele hits. Dat zijn links naar webpagina's waarop je ingetikte zoektermen voorkomen. Je hoeft dan maar zo'n link aan te klikken om de bijbehorende pagina in je browservenster op te roepen. Via de link VOLGENDE (en VORIGE) kan je door de lijst van hits bladeren... Maar welke magie schuilt nu achter zo'n supersnelle speurtocht?



De zoekknop leidt je standaard naar MSN Web Search.

Robots en gidsen

Je moet weten dat je achter de schermen met een zogenaamde zoekrobot wordt verbonden, en in dit geval is dat MSN Web Search. Dat is niets meer of minder dan een website waaraan een gegevensbank is gekoppeld. Wie niet vies is van enige vereenvoudiging, mag zich zo'n databank voorstellen in twee kolommen: links een ellenlange lijst met typische zoektermen, rechts telkens de *url's* van webpagina's waarop die zoektermen voorkomen. Een klein voorbeeldje ter verduidelijking:

Britney	url1	url5	url6	url10
Spears	url3	url5	url10	url14

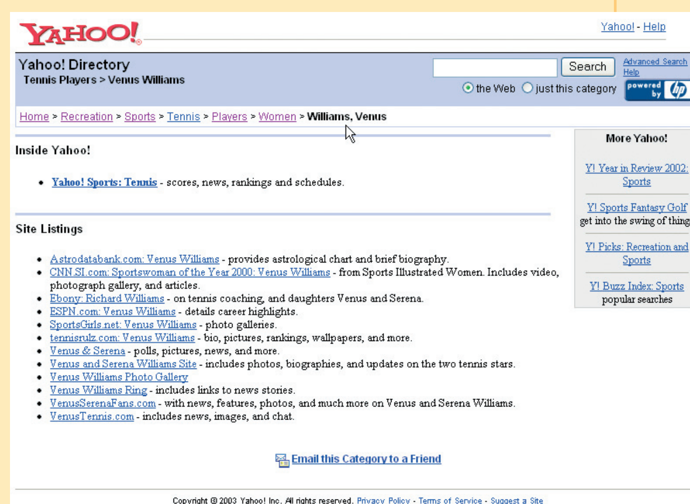
Stel dat je als zoekterm de woordcombinatie 'Britney Spears' intikt, dan zou deze databank je twee hits opleveren: url5 en url10. Je kan je nu best voorstellen dat zo'n gegevensbank reusachtig groot is en de vraag rijst dan ook: wie houdt die netjes up-to-date? Dat is zo'n titanenarbeid dat hiervoor de hulp van speciale software wordt ingeroepen. Kleine programmaatjes, zoek(ro)bots genoemd, speuren met een ontstellende snelheid het hele web af, op zoek naar pagina's, springend van link naar link. Tegelijk stockeren ze in hun databanken zowat alle woorden die ze op die pagina's inlezen, inclusief een link naar het webadres van die pagina's. Krachtige zoekrobots slagen er op die manier in wel miljoenen pagina's per dag te verwerken. Bekende, internationale voorbeelden zijn: AlltheWeb [**www.alltheweb.com**], AltaVista [**www.altavista.com**], Excite [**www.excite.com**], Google [**www.google.com**], HotBot [**www.hotbot.com**], Lycos [**www.lycos.com**], MSN [**www.msn.be**] en Northern

Yahoo! Moeder van alle webgidsen



Een webgids als Yahoo! laat je ook thematisch zoeken.

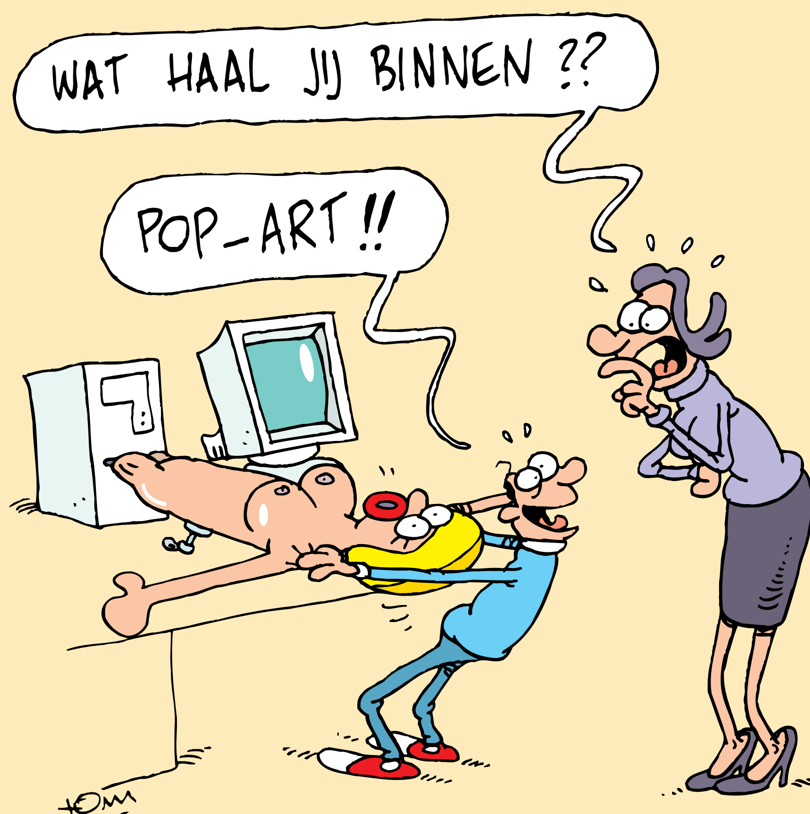
Om dat te begrijpen, moet je eerst weten hoe een typische webgids in elkaar steekt. Zo'n gids is eigenlijk een collectie van verschillende categorieën die elk een ander thema behandelen. Deze verzameling is hiërarchisch opgebouwd. Dat betekent dat een basiscategorie (zoals Recreatie) verschillende subcategorieën (zoals Hobby's, Sport en Reizen) kan bevatten. Die kunnen op hun beurt ook weer kleinere categorieën bevatten (zoals Boksen, Voetbal en Tennis), enz. Zo'n doordachte, thematische opbouw kan een flinke opsteker voor je speurtocht betekenen... Stel even dat je op zoek bent naar informatie over Venus Williams, maar je herinnert je enkel nog haar voornaam. Dan kan je als volgt te werk gaan in een webgids als Yahoo!. Tik VENUS in op de zoekbalk, en klik op de pagina die Yahoo! je vervolgens aanbiedt, de Category match 'TENNIS PLAYERS > VENUS WILLIAMS' aan. Herinner je je zelfs haar voornaam niet meer, dan kan je ook eigenhandig afdalen in de hiërarchische structuur, door achtereenvolgens deze categorieën aan te klikken: RECREATION > SPORTS > TENNIS > PLAYERS > WOMEN. In beide gevallen beland je overigens op dezelfde pagina met een uitgelezen selectie links naar documenten rond sportster Venus Williams. Bijkomend pluspunt is dat Yahoo! je onderweg ook links toont naar thematisch verwante onderwerpen, in dit geval andere tennissterren als Anna Kournikova, Chris Evert, Justine Henin-Hardenne, enz.



Hoe je ook zoekt, het resultaat blijft hetzelfde.

n

Light [**www.northernlight.com**]. Zoekrobots die hun zoekterrein tot België (kunnen) beperken zijn er ook, zoals: [**www.skynet.be**], [**www.webwatch.be**] en [**www.ilse.be**]. Zoekrobots moeten het natuurlijk van hun snelheid en dus van hun indrukwekkende databanken hebben, maar erg slim zijn ze niet bepaald. Tik je bijvoorbeeld 'Venus' in als zoekterm, dan mag je je aan een bonte verzameling van sportsites (Venus Williams) en wetenschappelijke sites (planeet Venus) verwachten... Je kan dan natuurlijk zelf je gezond verstand gebruiken en de zoekopdracht door bijkomende zoektermen verder afbakenen (genre: Venus Williams tennis), maar je kan het ook over een andere boeg gooien... Er bestaan namelijk ook databanken die niet door bots, maar effectief door mensen zijn samengesteld en die je eveneens via een website kan bevragen. Die noemen we wel eens webdirectories of webgidsen, en bekende voorbeelden zijn alvast Open Directory [**http://dmoz.org**] en Yahoo! [**www.yahoo.com**]. Voor Belgische sites kan je onder meer terecht bij [**www.in.be/nl/webs**], [**www.hotbel.be**] en [**http://yellow.advalvas.be/weblist.html**].
Intussen bestaan er trouwens al hybride vormen: zoekrobots dus die zelf ook dankbaar putten uit de databanken van zo'n webgids. Typisch voorbeeld is [**http://directory.google.com**] dat z'n mosterd haalt uit het Open Directory project. Maar welk voordeel kunnen zulke webgidsen nu hebben ten opzichte van zoekrobots?



Google: de populairste zoekrobot



Google: bedrieglijk simpel.

Ook al kunnen uitgebalanceerde webgidsen als Yahoo! je dus onmiskenbare voordelen bieden, ze kunnen niet echt tippen aan de populariteit van zoekrobots – Google op kop! Het principe van zo'n zoekrobot is overigens eenvoudig genoeg: je tikt één of meer zoekwoorden in de balk en je drukt op je **Entertoets**. Even later krijg je dan de hitlijst te zien waar je de gewenste links maar hoeft aan te klikken. Google zelf blinkt overigens uit door eenvoud en een handvol basistips zetten je al een heel eind op weg:

- Tik je meer dan één woord in, dan zal Google enkel die pagina's tonen die ál je ingetikte woorden bevatten. Wil je dat Google ook met pagina's aandraaft waarop slechts één van je ingetikte woorden voorkomt, dan moet je het woordje **OR** (in hoofdletters!) tussen die woorden inlassen (zoals in: shuttle discovery OR columbia).
- Google maakt géén onderscheid tussen hoofdletters en kleine letters (zoekterm Brussel levert dezelfde hits op als brussel of BRUSSEL).
- Wil je verschillende opeenvolgende woorden als één begrip laten beschouwen, plaats ze dan tussen aanhalingstekens (zoals in: "clickx magazine"). Google toont dan enkel die pagina's waarin die woorden onmiddellijk na elkaar voorkomen.
- Google negeert typische stopwoorden, losse letters en cijfers. Wil je dat Google daar toch rekening mee houdt, zet er dan een plusteken voor (zoals in: "Karel +I").
- Wil je absoluut vermijden dat bepaalde woorden in de gevonden webpagina's voorkomen, zet er dan een minteken voor (zoals in: venus -planet).
- Google kan niet overweg met deels ingetikte woorden, ook niet als je jokertekens gebruikt (mag of mag* zal je dus geen pagina's met magazine of magazines opleveren).

Geavanceerd zoeken in Google



Heel wat paardenkracht onder Google's motorkap!

Wie denkt dat daarmee alles is gezegd over Google, zit op het verkeerde spoor! Druk maar even de link **Geavanceerd zoeken** in op de hoofdpagina van Google. Het aantal opties waarmee je hier je zoekopdrachten kan verfijnen, is indrukwekkend. Een selectief overzicht...

In het blauwe kadertje maak je duidelijk hoe Google precies met je ingetikte zoektermen moet omspringen. Liever dan je het hoofd te breken over aanhalingstekens, OR's, plusjes of minnetjes vul je hier je zoektermen gewoon in de juiste balk in. Tik je bijvoorbeeld 'Clickx magazine' in bij **Met de exacte woordcombinatie** dan wordt dat als één begrip beschouwd, maar tik je deze woorden in bij **Met één van deze woorden**, dan laat Google je niet enkel de pagina's zien die 'Clickx magazine' bevatten, maar ook die met enkel 'clickx' of 'magazine'. Je kan je zoektocht hier echter ook nog op totaal andere manieren beperken. Heel leuk is bijvoorbeeld dat je het gewenste bestandsformaat kan aanstippen (waaronder .doc, .pdf, .ppt en .xls). In dat geval krijg je niet langer gewone webpagina's te zien, maar beperkt Google zich tot het bestandsformaat dat je hebt aangestipt. Zo kan je bijvoorbeeld heel snel op zoek gaan naar kant-en-klare PowerPoint-presentaties (.ppt)! Verder kan je ook nog aangeven dat je enkel in min of meer recente webpagina's bent geïnteresseerd. Hoever je precies in de tijd wil teruggaan, geef je dan bij de rubriek **Datum** te kennen. Je bent hier jammer genoeg wel beperkt tot 3 maanden, 6 maanden of één jaar. Bij **Taal** kan je ook nog instellen dat je enkel interesse hebt in pagina's geschreven in een welbepaalde taal (zoals Nederlands). Stel even dat het Nederlands van onze noorderburen je niet bepaald kan charmeren: hoe filter je die pagina's er dan uit? Ook dát kan... bij **Domein** hoef je dan maar aan te duiden: **Geen** resultaten weergeven van de site of het domein **.nl** (.nl is namelijk de landcode van sites uit Nederland). Wil je een paar talen tegelijk selecteren, dan moet je wel elders zijn. Daarvoor moet je op de hoofdpagina van Google de link **Voorkeuren** aanklikken. Bij **Zoektaal** stip je vervolgens de optie **Enkel zoeken in de geselecteerde talen aan**, en zet je een vinkje voor alle talen die je interesseren. Druk vervolgens op de knop **Voorkeuren opslaan**: Google onthoudt je voorkeuren dan, en zal daar ook bij je volgende zoektochten automatisch rekening mee houden – op voorwaarde natuurlijk dat je dezelfde computer gebruikt.



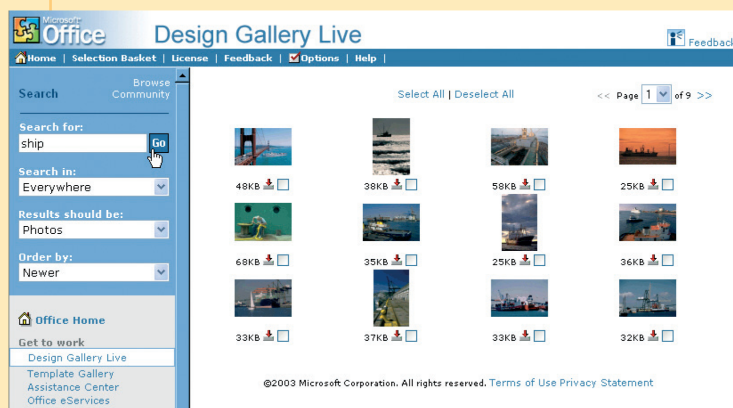
Leg je voorkeuren in Google vast.

Méér dan webpagina's

Zoekrobots verslinden dus massa's webpagina's, maar soms willen we best wat kieskeuriger zijn. Wat als je bijvoorbeeld enkel in leuke plaatjes geïnteresseerd bent? Ook dan kan je prima terecht bij Google: op de hoofdpagina moet je dan wel eerst het tabblad **AFBEELDINGEN** aanklikken. Voor de rest gebeurt het zoeken op dezelfde manier als anders, inclusief de pagina met geavanceerde zoekopties! Ook de Microsoft Design Gallery bevat een imposante verzameling illustraties, tot zelfs webanimaties en geluiden toe [**http://dgl.microsoft.com**]. Deze mediagalerie kan je overigens ook rechtstreeks bereiken vanuit een Office-programma als MS Word (**INVOEGEN**, **FIGUUR**, **ILLUSTRATIE**). Tip: kies op de site via **OPTIONS** voor de taalversie **ENGELS**: je krijgt dan veel meer illustraties te zien, op voorwaarde weliswaar dat je dan Engelstalige zoektermen intikt.
Ben je van het nieuws-gierige type en ben je dagelijks naarstig op zoek naar interessante nieuwsflashes, dan kan je vooral terecht op [**http://news.google.com**] en op AlltheWeb (tabblad **NEWS**). Ook het vernieuwde AltaVista heeft recent een nieuwsrubriek (tabblad **NEWS**) toegevoegd. Bij deze laatste kan je trouwens ook aankloppen voor een niet onaardige verzameling mp3/Audio- en Video-links.



Met Google haal je er snel de afbeeldingen tussenuit.



Op zoek naar leuke beelden en animaties in de Design Gallery.

Altijd bij de hand

Wedden dat je in minder dan geen tijd uitgroeit tot een fervente speurneus? In dat geval wordt het natuurlijk al snel omslachtig telkens een ommetje naar een van je favoriete zoeksites te moeten maken. Gelukkig laten heel wat zoeksites je toe gratis een handige zoekbalk te installeren, zodat je rechtstreeks vanuit je browser een zoekopdracht kan intikken. Populaire exemplaren zijn alvast die van Google [**http://toolbar.google.com**], AskJeeves [**http://sp.ask.com/docs/toolbar**] en Yahoo! [**http://companion.yahoo.com**].
Nadeel is wel dat zo'n zoekbalk je beperkt tot die bepaalde zoeksite. Dat is echter niet het geval met nieuwkomer UltraBar [**www.ultrabar.com**]. Die heeft namelijk een lijstje met zo'n twintig zoekmachines aan boord en je kan zelf aanstippen welke je erbij wil. Het is zelfs mogelijk een eigen zoekmachine die niet in het lijstje is voorzien, toe te voegen. Enigszins vergelijkbaar hiermee is de krachtige Dave's Quick Search Deskbar [**http://notesbydave.com/toolbar/doc.htm**], die zich in de taakbalk van Windows nestelt.



Met de zoekbalk van Google heb je je favoriete zoekmachine altijd bij de hand!

Leuke uitsmijter...

Zit je zonder browser, maar heb je wel zo'n sjieke *PDA* of *WAP*-machine bij de hand, geen nood: ook dán kan je nog zoekopdrachten doorsturen! Je hoeft enkel een mailtje te sturen naar [**google@capeclear.com**] met je zoekterm in de onderwerpregel en een paar seconden later rolt de hitlijst al in je mailbox. Pik daar vervolgens de leukste url uit, tik die op zijn beurt in de onderwerpregel van een nieuw berichtje in, dat je dan naar [**page@grabpage.org**] stuurt. Even later krijg je dan alle tekst van die webpagina in een mailtje te lezen...

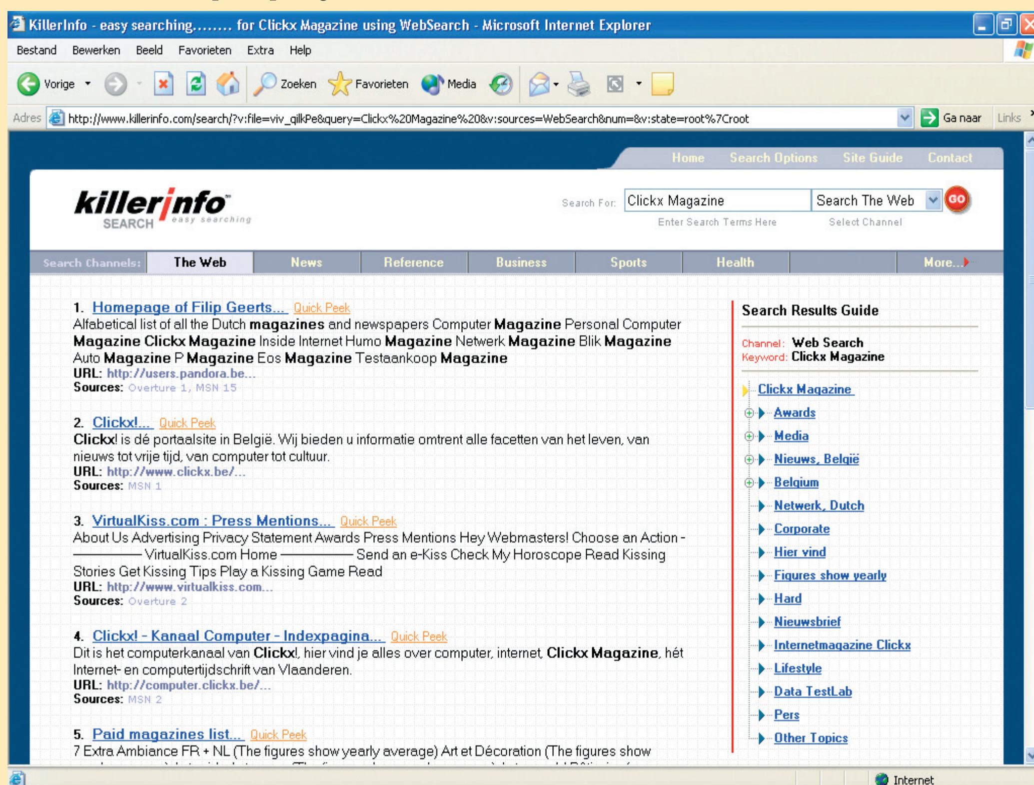
Zoeken in de zoekers

Het probleem met al die zoekmachines is dat ze niet alleen vaak massa's (onbruikbare) hits uitspuwen, maar dat ze stuk voor stuk ook hun sterke en zwakke kanten hebben. Het zou dus best nuttig zijn als je vanuit één programma verschillende zoekmachines tegelijk kan doorzoeken, waarna je een gefilterd en geordend overzicht krijgt van de beste hits... Zulke programma's bestaan, en ze hebben ook een naam: metasearchers! Bekende voorbeelden zijn Search.com [www.search.com] en vooral ook MetaCrawler [www.metacrawler.com]. Via de link GEAVANCEERD ZOEKEN kan je overigens zelf bepalen aan welke zoekmachines ze jouw zoekterm moeten doorgeven. Een jonge loot aan de stam van de metasearchers is KillerInfo [www.killerinfo.com]. Bijzonder leuk hieraan is de Quick Peek, waarmee je in een minivenster al een eerste blik kan werpen op de gevonden hits. KillerInfo onderneemt ook waardige pogingen om alle hits per thema te groeperen, zodat je nog gerichter kan zoeken. Sommige metasearchers komen echter in de vorm van een downloadbaar programma, dat je dus zelf eerst moet installeren. Heel populair is alvast Copernic Agent, dat je in een basisversie gratis kan afhalen op [www.copernic.com]. Een iets minder veelzijdig alternatief is WebFerret 5 dat je gratis binnenhaalt op [www.ferretsoft.com]. Verder is er nog het relatief onbekende Glooton [www.glooton.com] dat nochtans wel de moeite loont: via een systeem van in- en uitschakelbare plug-ins kan je het programma namelijk aanpassen aan je eigen behoeften.

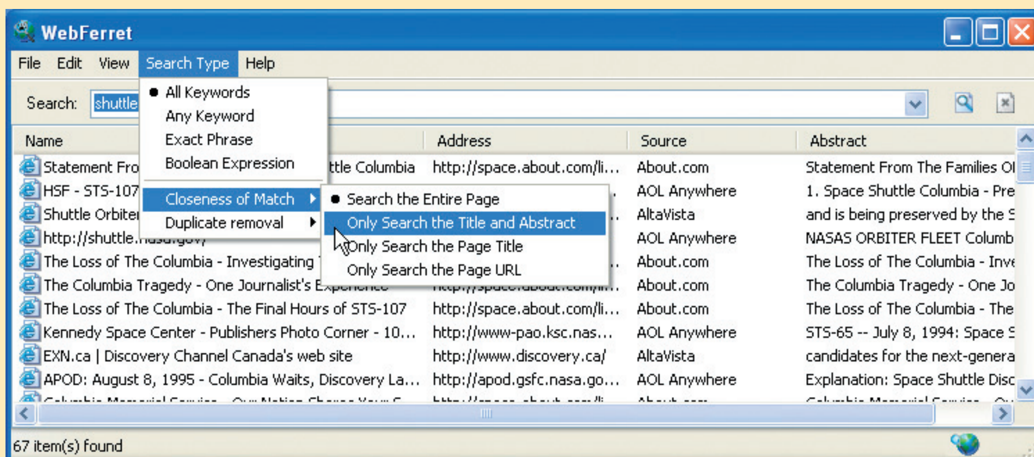
Via de link VOORKEUREN kan je een zoekmachine als Google toch al een beetje naar je hand zetten. Ook andere zoekrobots – zoals het onlangs vernieuwde HotBot – laten je toe de zoekinterface te personaliseren. Maar je kan gerust nog een stapje verder gaan, en een volledige webpagina naar eigen smaak inrichten en die vervolgens als startpagina van je browser instellen. Zo kan je bij heel wat grote zoeksites (zoekportalen genoemd) als Yahoo! [http://my.yahoo.com] en Excite [http://my.excite.com] niet alleen het uitzicht, maar ook de inhoud van de hoofdpagina zelf bepalen. Ook nieuwkomer My Way [www.myway. com] laat zich niet onbetuigd: niet geplaagd door enige reclameboodschappen laat deze site zich namelijk heel gewillig naar jouw wensen kneden. De zoekmotor van My Way mag er overigens ook zijn, want die is blijkbaar 'powered by Google'!



De nieuwe HotBot laat zich sterk personaliseren.



KillerInfo: een jonge metasearcher.



Webferret: iets minder veelzijdig, maar wel gratis.

— Toon Van Daele —